数理解析研究所講究録
70巻 1969年 47-66

Some Problems on Non-linear Stochastic Control Processes*

by

Yoshifumi Sunahara**

1. Introduction and symbolic conventions

During the past decade, the problem of finding the optimal control has received a great deal of interest as results of the ever-complicated demand to controls and ever-increasing complexity of the operation of modern systems. However, most of this work has concentrated on completely linear dynamical systems, neglecting the effects of nonlinear characteristics exhibited in practice. There is no needs to say that dynamical systems to be controlled exhibit various kinds of nonlinear characteristics and may operate in a random environment whose stochastic characteristics undergo drastic changes. Thus, the general problem to solved is to find the control of a noisy nonlinear dynamical system in some optimal fashion, given only partial and noisy observations of system state and, possibly, only an incomplete knowledge of the system. Under such conditions as linearity of the dynamical system, noisy observation and performance criterion given by a quadratic cost functional, it has already been shown that the optimal control problem and the optimal estimation problem of the system state from the noise-corrupted observations may independently be solved. [1]∿[3] However, this is not the case in general for the optimal control of nonlinear dynamical systems, and the overall problem of optimal control and estimation must be carried out simultaneously. Since the establishment of

---

* December 23, 1968. The materials of this short note have been taken from the author's papers submitted to the Division of Automatic Control, ASME, the International Journal on Control and the Jorunal of the Franklin Institute.

** Kyoto Institute of Technology, Matsugasaki, Kyoto, Japan.

the precise technique for the state estimation and the optimal control of nonlinear dynamical systems is almost impossible, in this paper, the author will introduce the reader to an approximate method which will be shown to play an important role in the realization of a broad class of stochastic optimal control and which will hopefully be of an extensive use to the version of computing control in industrial processes.

Throughout this paper, we use the same symbols for the true stochastic processes and for the quasi-linear stochastic processes which are the approximations to the true one by the method described later. Vector and matrix notations follow the usual manner, that is, lower case letters a, b and c, ··· will denote column vectors with i-th real components a_i, b_i, and c_i, ···. Capital letters A, B, C and G, ··· denote matrices with elements a_{ij}, b_{ij}, c_{ij} and g_{ij}, ··· respectively. If M is a matrix, then M' denotes its transpose. The symbol $|M|$ denotes the determinant of the matrix M.

Certain algebraic quantities such as algebras, fields, etc., are expressed by the symbols, $\mathcal{Y}$, $\mathcal{F}$, ···, etc. The symbol, $\mathcal{Y}_t$, denotes the smallest σ-algebra of ω sets with respect to which the random variables $y(\tau)$ with $\tau \leq t$ are measurable, where ω is the generic point of the probability space Ω. The mathematical expectation is denoted by E. The conditional expectation of a random variable conditioned by $\mathcal{Y}_t$ is simply expressed by "$\hat{\cdot}$" such that $E\{x(t)|\mathcal{Y}_\tau\} = \hat{x}(t|\tau)$, where $\tau \leq t$.

For convenience of the present description, the principal symbols used here are listed below:

t: time variable, particularly the present time

t_0: the in tial time at which observations start

$x(t)$ and $y(\tau)$: n-dimensional vector stochastic processes representing the system states and the observations respectively.

$u(t)$: m-dimensional control vector taking values in a convex compact subset $U \subset E^m$ where E^m is the m-dimensional Euclidean space

$w(t)$ and $v(t)$: d_1- and d_2-dimensional Brownian motion processes respectively

$C(t)$, $G(t)$ and $R(t)$: $n \times m$, $n \times d_1$ and $n \times d_2$ matrices whose components depend on t

$f[t, x(t)]$ and $h[t, x(t)]$: n-dimensional vector valued nonlinear functions respectively

$\hat{x}(t|t)$: optimal estimate of $x(t)$ conditioned by $\mathcal{Y}_t$, i.e., $E\{x(t)|\mathcal{Y}_t\} \triangleq \hat{x}(t|t)$

$P(t|t)$: an error covariance matrix in optimal estimate of $x(t)$ conditioned by $\mathcal{Y}_t$, i.e., $P(t|t) \triangleq \text{cov.}\,[x(t)|\mathcal{Y}_t]$.

2. Mathematical models and problem statement

Guided by a well-known state space representation concept, the dynamics of an important class of dynamical systems can be described by a nonlinear vector differential equation,

$$\frac{dx(t,\omega)}{dt} = f[t, x(t,\omega)] - C(t)u(t) + G(t)\gamma(t,\omega), \tag{2.1}$$

where $\gamma(t, \omega)$ is a d_1-dimensional Gaussian white noise disturbance. For the economy of descriptions, we shall omit to write the symbol ω in the sequel because of no confusion.

We shall start with a precise version of Eq. (2.1), namely the stochastic differential equation of Itô-type, [4]

$$dx(t) = f[t, x(t)]dt - C(t)u(t)dt + G(t)dw(t), \tag{2.2}$$

where the d_1-dimensional Brownian motion process $w(t)$ has been introduced here along the relation between a Brownian motion process and a white noise or a sufficiently wide (but finite) band Gaussian random process $\gamma(t)$, [5]~[6]

(2.3) $$w(t) \stackrel{\sim}{=} \int^{t} \gamma(s)ds.$$

We suppose that observations are made at the output of the nonlinear system with additive Gaussian disturbance. The observation process $y(t)$ is the n-dimensional vector random process determined by

(2.4) $$dy(t) = h[t, x(t)]dt + R(t)dv(t),$$

where we assume that the system noise $w(t)$ and the observation noise $v(t)$ are mutually independent.

In practical terms, the problem is to control $x(t)$ in such a way as to minimize a real valued functional,

(2.5) $$J(u) = E\{\int_{t_0}^{T} L[t, x(t), u(t)]dt\}, \quad [t_0 \leq t \leq T]$$

based on the a priori probability distribution of $x(t_0)$, provided that the process $y(s)$ for $t_0 \leq s \leq t$ is acquired as the observation process, where $y(t_0) = 0$ and where L and L_u are bounded, uniformly Hölder continuous in t and uniformly Lipschitz continuous in x, and where L_{uu} is bounded and continuous on $[t_0, T] \times E^n \times U$. The subscript denotes differentiation here and below.

We shall consider the case where the state variables $x(t)$ are completely observable. Usually, in this case, the optimal control must be assumed to depend on $x(s)$, where $t_0 \leq s \leq t$. Bearing this fact in mind, and following [3], we shall proceed to establish the solution of the stochastic differential equation (2.2).

Let G denotes the class of continuous functions $\lambda(t)$ which are defined on $[t_0, T]$ and which take the values in E^n, and F_t denotes a functional operator in E^n. Clearly, if $\lambda \varepsilon G$, then $F_t \lambda \varepsilon G$. Furthermore, let ψ denotes a mapping of $[t_0, T] \times G$ onto U with the following properties:

P-1: For each $\lambda \varepsilon G$, the functional $\psi(t, \lambda)$ is Hölder continuous with respect to t.

P-2: For $t\varepsilon[t_0, T]$, the functional ψ satisfies a uniform Lipschitz condition, i.e.,

$$\|\psi(t, \lambda) - \psi(t, \lambda_0)\| \leq K_1 \|\lambda - \lambda_0\|_{\sup},$$

where the function $\lambda_0 \varepsilon G$ and K_1 is a real positive constant, and where $\|\cdot\|_{\sup}$ expresses sup. norm in G.

Let $\Psi(t, \cdot)$ be an m-dimensional vector stochastic process, such that, for each $t\varepsilon[t_0, T]$, $\psi(t, \cdot)$ is measurable and

$$\int_{t_0}^{T} E\{\|\psi(t, \cdot)\|^2\}dt<\infty, \tag{2.6}$$

where $\|\cdot\|$ expresses the norm in E^m. Let Ψ be the class of the $\psi(t, \cdot)$-process. For some $\psi\varepsilon\Psi$, we call the u(t) admissible and write $u\varepsilon U$, if

$$u(t) = \psi(t, \cdot) \qquad \text{for } t\varepsilon[t_0, T]. \tag{2.7}$$

For the security of mathematical development in the sequel, the following hypotheses are additionally made:

H-1: The component of the functions $f[\cdot, \cdot]$ and $h[\cdot, \cdot]$ are Baire functions with respect to the pair (t, ξ) for $t_0 \leq t \leq T$ and $-\infty<\xi<\infty$, where $x(t) = \xi$.

H-2: The functions $f[\cdot, \cdot]$ and $h[\cdot, \cdot]$ satisfy a uniform Lipschitz conditions in the variable ξ and are bounded respectively by

$$\|f(t, \xi)\| \leq K_2(1 + \xi'\xi)^{1/2} \tag{2.8a}$$

and

$$\|h(t, \xi)\| \leq K_3(1 + \xi'\xi)^{1/2} \tag{2.8b}$$

where both K_2 and K_3 are real positive constants and are independent of both t and ξ respectively.

H-3: The functions $f[\cdot, \cdot]$ and $g[\cdot, \cdot]$ are uniformly Hölder continuous in t with exponent α.

H-4: $x(t_0)$ is a random variable independent of the $w(t)$-process.

H-5: All parameter matrices are measurable and bounded on the finite time interval $[t_0, T]$.

H-6: $\{R(t)R(t)'\}^{-1}$ exists and this is bounded on $[t_0, T]$.

With the properties P-1 and P-2 and the hypotheses H-1 to H-6, Eq. (2.2) has exactly a unique continuous solution $x(t)$. A precise interpretation of Eq. (2.2) is given by the stochastic integral equation of Itô-type [4]:

$$x(t) = x(t_0) + \int_{t_0}^{t} f[s, x(s)]ds - \int_{t_0}^{t} C(s)u(s)ds + \int_{t_0}^{t} G(s)dw(s). \tag{2.9}$$

3. Quasi-linear stochastic differentials and an approximation to non-linear filtering equations

In this section, the development of the discussion requires that, until further notice, we set the control $u(t)$ equals to zero in Eq. (2.2). When $u(t) = 0$, the symbol is temporarily changed from $x(t)$ to $z(t)$. With this symbolic change, Eq. (2.2) is

$$dz(t) = f[t, z(t)]dt + G(t)dw(t) \tag{3.1}$$

and also Eq. (2.4) is written by

$$dy(t) = h[t, z(t)]dt + R(t)dv(t), \tag{3.2}$$

where the same symbol $y(t)$ has been used as in Eq. (2.4) because of economy of notations.

The problem considered here is to find the minimal variance estimate of the state variable $z(t)$, provided that the process $y(s)$ for $t_0 \leq s \leq t$ is acquired as the observation process, where $y(t_0) = 0$.

We expand the function, $f(t, z)$, in Eq. (3.1) into

$$f[t, z(t)] = a(t) + B(t)\{z(t) - \hat{z}(t|t)\} + e(t), \tag{3.3}$$

where $a(t)$, $B(t)$ are an n-dimensional vector and an $n\times n$ matrix respectively, and where $e(t)$ denotes the collection of n-dimensional vector error terms, and where $\hat{z}(t|t) = E\{z(t)|\mathcal{Y}_t\}$. We shall determine $a(t)$ and $B(t)$ in such a way that the conditional expectation of the squared norm of $e(t)$ conditioned by $\mathcal{Y}_t$, $E\{\|e(t)\|^2|\mathcal{Y}_t\}$, becomes minimal with respect to $a(t)$ and $B(t)$. It is a simple exercise to show in the calculus of variation that the necessary and sufficient conditions for min. $E\{\|e(t)\|^2|\mathcal{Y}_t\}$ are given by

$$a(t) = E\{f[t, z(t)]|\mathcal{Y}_t\} \tag{3.4a}$$

and

$$B(t) = E[\{f[t,z(t)] - \hat{f}[t,z(t)]\}\{z(t) - \hat{z}(t|t)\}'|\mathcal{Y}_t]\bar{P}(t|t)^{-1}, \tag{3.4b}$$

where

$$\bar{P}(t|t) = \text{cov. } [z(t)|\mathcal{Y}_t] \tag{3.5}$$

The scalar expressions of (3.4) are as follows:

$$a_i(t) = E\{f_i[t, z(t)]|\mathcal{Y}_t\} = \hat{f}_i[t, z(t)] \tag{3.6a}$$

$$\sum_{\nu=1}^{n} b_{i\nu}(t)E[\{z_\nu(t) - \hat{z}_\nu(t|t)\}\{z_j(t) - \hat{z}_j(t|t)\}|\mathcal{Y}_t]$$

$$= E[\{f_i[t, z(t)] - \hat{f}_i[t, z(t)]\}\{z_j(t) - \hat{z}_j(t|t)|\mathcal{Y}_t\} \tag{3.6b}$$

where $\hat{z}_j(t|t) = E\{z_j(t)|\mathcal{Y}_t\}$ and $i, j=1,2, \cdots, n$. Using $a(t)$ and $B(t)$ determined by (3.4) and (3.5), we approximate Eq. (3.1) by

$$z(t) = z(t_0) + \int_{t_0}^{t} [a(s) + B(s)\{z(s) - \hat{z}(s|s)\}]ds + \int_{t_0}^{t} G(s)dw(s). \tag{3.7}$$

The same procedure is applicable to the observation process given by Eq. (3.2). Through the expansion of the function, $h[t, z(t)]$, in the form;

(3.8) $\quad h[t, z(t)] = h_1(t) + H_2(t)\{z(t) - \hat{z}(t|t)\} + e_h(t),$

the following conditions can easily be obtained so as to minimize $E\{\|e_h(t)\|^2|\mathcal{Y}_t\}$ with respect to $h_1(t)$ and $H_2(t)$:

(3.9a) $\quad h_1(t) = E\{h[t, z(t)]|\mathcal{Y}_t\} \triangleq \hat{h}[t, z(t)]$

(3.9b) $\quad H_2(t) = E[\{h[t,z(t)] - \hat{h}[t,z(t)]\}\{z(t) - \hat{z}(t|t)\}'|\mathcal{Y}_t]\bar{P}(t|t)^{-1}.$

We shall assume here that, for $t\varepsilon[t_0, T]$, the conditional probability density function $p\{z(t)|\mathcal{Y}_t\}$ is Gaussian with the mean value $\hat{z}(t|t)$ and covariance matrix $\bar{P}(t|t)$, i.e.,

(3.10) $$p\{z(t)|\mathcal{Y}_t\} = (2\pi)^{-\frac{n}{2}}|P(t|t)|^{-\frac{1}{2}}\exp[-\frac{1}{2}\{z - \hat{z}(t|t)\}'\bar{P}(t|t)^{-1} \times \{z - \hat{z}(t|t)\}].$$

With the help of (3.10), both $a(t)$ and $B(t)$ can be obtained in the form $a(t) = a(t, \hat{z}(t|t), \bar{P}(t|t))$ and $B(t) = B(t, \hat{z}(t|t), \bar{P}(t|t))$ or $b_{ij}(t) = \partial a_i(t)/\partial \hat{z}_j(t|t)$. A striking fact is that the random variables $a(t)$ and $B(t)$ are not independent but depend mutually on the state estimate $\hat{z}(t|t)$ and the error covariance matrix $\bar{P}(t|t)$. From this point of view, in reality, more precise symbols, $a(t, \hat{z}(t|t), \bar{P}(t|t))$ and $B(t, \hat{z}(t|t), \bar{P}(t|t))$ should be introduced. However, for the economy of description, we merely denote these by $a(t)$ and $B(t)$ without indicating the dependence on both $\hat{z}(t|t)$ and $\bar{P}(t|t)$. Both $h_1(t)$ and $H_2(t)$ also follow this symbolic convention.

From Eq. (3.7), we may thus define here the following n-dimensional quasi-linear stochastic differentials of Itô-type for Eq. (3.1),

(3.11) $\quad dz(t) = B(t)z(t)dt + \{a(t) - B(t)\hat{z}(t|t)\}dt + G(t)dw(t),$

and for the observation process (3.2),

(3.12) $\quad dy(t) = H_2(t)z(t)dt + \{h_1(t) - H_2(t)\hat{z}(t|t)\}dt + R(t)dv(t).$

However, respective drift terms in Eqs. (3.11) and (3.12) still remain unknown. We shall thus proceed to solve the problem including the computation of the state estimate $\hat{z}(t|t)$ and the error covariance matrix $\bar{P}(t|t)$.

Let $\Phi(t, t_0)$ be the fundamental matrix associated with the homogeneous differential equation, $dz(t)/dt = B(t)z(t)$. The solution of Eq. (3.11) can formally be written as

$$(3.13) \qquad z(t) = \Phi(t, t_0)z(t_0) + \int_{t_0}^{t} \Phi(t, s)\{a(s) - B(s)\hat{z}(s|s)\} ds + \int_{t_0}^{t} \Phi(t, s)G(s)dw(s).$$

We write for the second term of the right side of Eq. (3.13)

$$(3.14) \qquad \zeta(t) = -\int_{t_0}^{t} \Phi(t, s)\{a(s) - B(s)\hat{z}(s|s)\}ds$$

and introduce a new stochastic process

$$(3.15) \qquad \xi(t) = z(t) + \zeta(t).$$

Combining Eq. (3.13) with (3.14) and noting that $\xi(t_0) = z(t_0)$, from Eqs. (3.14) and (3.15), the $\xi(t)$-process is of Itô-type and the stochastic differential is

$$(3.16) \qquad d\xi(t) = B(t)\xi(t)dt + G(t)dw(t).$$

On the other hand, it follows from Eq. (3.12) that

$$(3.17) \qquad y(t) = \int_{t_0}^{t} H_2(s)z(s)ds + \int_{t_0}^{t} \{h_1(s) - H_2(s)\hat{z}(s|s)\} ds + \int_{t_0}^{t} R(s)dv(s).$$

Let the second term of the right side of Eq. (3.17) be $\zeta_y(t)$ and define $\eta_y(t) \triangleq y(t) - \zeta_y(t)$. Then we obtain

$$(3.18) \qquad d\eta_y(t) = H_2(t)z(t)dt + R(t)dv(t)$$

with $\eta_y(t_0) = 0$. With $\eta_y(t)$ determined by Eq. (3.18), define a new stochastic process $\eta(t)$ by its stochastic differential,

(3.19) $$d\eta(t) = d\eta_y(t) + H_2(t)\zeta(t)dt,$$

and $\eta(t_0) = 0$. Using Eqs. (3.15) and (3.18), Eq. (3.19) becomes

(3.20) $$d\eta(t) = H_2(t)\xi(t)dt + R(t)dv(t).$$

Since $\zeta(t)$ is $\mathcal{Y}_t$-measurable, it follows from Eq. (3.15) that

(3.21) $$\hat{\xi}(t|t) = E\{\xi(t)|\mathcal{Y}_t\} = \hat{z}(t|t) + \zeta(t).$$

Let $\mathcal{H}_t$ be the σ-algebra of ω sets generated by the random variables $\eta(s)$ for $t_0 \leq s \leq t$. Then the y(t)-process is $\mathcal{H}_t$-measurable and thus

(3.22) $$E\{\xi(t)|\mathcal{Y}_t\} = E\{\xi(t)|\mathcal{H}_t\} \triangleq \hat{\xi}(t|t).$$

Now we consider that the $\xi(t)$-process is the fictitious state variables determined by Eq. (3.16) and that Eq. (3.20) denotes the observations which are made on the $\xi(t)$-process. This situation implies that the current estimate $\hat{\xi}(t|t)$ is given by [9], [10]

(3.23) $$d\hat{\xi} = B(t)\hat{\xi}dt + P_\xi(t|t)H_2(t)'\{R(t)R(t)'\}^{-1}\{d\eta - H_2(t)\hat{\xi}dt\},$$

where

(3.24) $$P_\xi(t|t) = \text{cov. } [\xi(t)|\mathcal{H}_t].$$

Substituting Eq. (3.20) into Eq. (3.23) and using Eqs. (3.12) and (3.21), it follows that

(3.25) $$d\hat{z} = \hat{f}[t, z(t)]dt + \bar{P}(t|t)H_2(t)'\{R(t)R(t)'\}^{-1}(dy - \hat{h}dt).$$

where Eqs. (3.6a) and (3.9a) have been used. By combining (3.21) with (3.24), we have

(3.26a) $$\bar{P}(t|t) = \text{cov. } [z(t)|\mathcal{Y}_t] = P_\xi(t|t)$$

and the version of $d\bar{P}(t|t)/dt$ is

(3.26b) $$\frac{d\bar{P}}{dt} = B\bar{P} + \bar{P}B' + GG' - \bar{P}H_2'\{RR'\}^{-1}H_2\bar{P}.$$

Equations (3.25) and (3.26) describe the dynamic structure of a quasi-linear filter for generating a current estimate $z(t|t)$ with the respectively given initial values, $\hat{z}(t_0|t_0)$ and $\bar{P}(t_0|t_0)$.

4. Quasi-optimal control

In this section, the control term $u(t)$ in Eq. (3.1) is revived, noting that the symbol changes naturally from $z(t)$ to $x(t)$.

Let the function L in (2.5) be

$$(4.1) \qquad L(t, x, u) = x'M(t)x + u'N(t)u,$$

where M and N are respectively measurable, locally bounded, positive semi-definite and positive definite symmetric matrices. In the case where both the dynamical system and the observation are respectively determined by linear stochastic differentials, it has already been varified that the optimal control exists and this is $u^o(t) = \psi^o[t, \hat{x}(t|t)] = N(t)^{-1}C(t)'Q(t)\hat{x}(t|t)$, where Q is the unique solution of a certain matrix Riccati equation. [1], [3] In the case of nonlinear regulator problems considered, the quasi-optimal control may be found out by an extensive use of the quasi-optimal control may be found out by an extensive use of the quasi linearization technique developed in the previous section to the version of stochastic control.

It is apparent that the $x(t)$-process has the quasi-linear stochastic differential,

$$(4.2) \qquad dx(t) = B(t)x(t)dt + \{a(t)-B(t)\hat{x}(t|t)\}dt - C(t)\psi(t)dt + G(t)dw(t),$$

where the definition of the admissible control given by (2.7) has been taken into account with the simplified notation $\psi(t.x)$. It also follows that

$$(4.3) \qquad dy(t) = h_1(t)dt + H_2(t)\{x(t) - \hat{x}(t|t)\}dt + R(t)dv(t).$$

Furthermore, with the help of Eq. (3.25), it can easily be shown that the state estimation $\hat{x}(t|t)$ for the nonlinear system described by Eq. (4.2) is

$$(4.4) \qquad d\hat{x} = \hat{f}dt - C\psi dt + PH_2'(RR')^{-1}(dy - \hat{h}dt),$$

where the version of dP/dt has the same form as given by Eq. (3.26b).

In the present case, the basic process is $\hat{x}(t|t)$ $(t_0 \leq t \leq T)$ with the stochastic differential (4.4); the cost rate function is given by (4.1) and the performance index by (2.5).

Combining the stochastic linearization technique with the line of attack on the linear regulator problem, we shall suppose that $u(t) = \hat{\psi}[t, \hat{x}(t|t)]$. It has been proved by solving the following Bellman's equation [10] that the optimal control $\hat{\psi}^o$ and $V(t, \xi)$ exist

$$(4.5a) \qquad \min_{u \in U}\{\hat{L}(t, \xi, u) + V_t(t, \xi) + \hat{\mathcal{L}}_{\hat{\psi}}V(t, \xi)\} = 0,$$

with terminal condition

$$(4.5b) \qquad V(T, \xi) = 0,$$

where

$$(4.6) \qquad V(t, \xi) = \mathcal{E}\{\textstyle\int_t^T \hat{L}[s, \hat{x}(t|t), \hat{\psi}^o[s, \hat{x}(s|s)]]ds | \hat{x}(t|t) = \xi\}$$

$$(4.7) \qquad \hat{L} = \mathcal{E}\{L[s, x(s), \hat{\psi}^o[s, \hat{x}(s|s)]] | \hat{x}(s|s) = \xi\}$$

and $\hat{\mathcal{L}}_{\hat{\psi}}$ denotes the differential generator of the $\hat{x}(t|t)$-process given by [11]

$$(4.8) \qquad \hat{\mathcal{L}}_{\hat{\psi}}(V) \equiv \tfrac{1}{2}\mathrm{tr}\{\textstyle\sum(t)'V_{\xi\xi}\sum(t)\} + \{a(t) - C(t)\hat{\psi}(t, \xi)\}'V_\xi$$

with

$$(4.9) \qquad \textstyle\sum(t) = PH_2'(RR')^{-1}R$$

because of (4.4) and the fact that the differential $dy - \hat{h}dt$ in Eq. (4.4) may be replaced by the suitably scaled differential of a Brownian motion process.

In the case where the function L is given by (4.1), it follows from (4.7) that

$$\hat{L}[s,\ \xi,\ \hat{\psi}(s,\ \xi)] = \mathrm{tr}\ M(s)P(s|s) + \xi'M(s)\xi + \hat{\psi}'N(s)\hat{\psi}. \tag{4.10}$$

We shall suppose that Bellman's equation (4.5) has a solution

$$V(t,\ \xi) = \xi'\Pi(t)\xi + 2\xi'\alpha(t) + \beta(t), \tag{4.11}$$

where $\Pi(t)$, $\alpha(t)$ and $\beta(t)$ will be determined as the solutions of matrix differential equations which will be given later. Applying (4.8), (4.10) and (4.11) to (4.5) and performing the minimization of Eq. (4.5), the optimal control is

$$\hat{\psi}^{o}(t,\ \xi) = \{N(t)^{-1}C(t)'\Pi(t)\}\xi + N(t)^{-1}C(t)'\alpha(t), \tag{4.12}$$

and $\Pi(t)$, $\alpha(t)$ satisfy

$$\frac{d\Pi(t)}{dt} - \Pi(t)C(t)N(t)^{-1}C(t)'\Pi(t) + M(t) = 0 \tag{4.13}$$

$$\frac{d\alpha(t)}{dt} - \Pi(t)C(t)N(t)^{-1}C(t)'\alpha(t) + \Pi(t)a(t) = 0 \tag{4.14}$$

for $t_0 \leq t \leq T$ with

$$\Pi(T) = 0,\ \alpha(T) = 0. \tag{4.15}$$

Furthermore, $\beta(t)$ in (4.11) satisfies

$$\frac{d\beta(t)}{dt} + \mathrm{tr}[\Sigma(t)'\Pi(t)\Sigma(t)] + \mathrm{tr}[M(t)P(t|t)] + 2a(t)'\alpha(t) - \alpha(t)'C(t)N(t)^{-1}C(t)'\alpha(t) = 0 \tag{4.16}$$

for $t_0 \leq t \leq T$ with

$$\beta(T) = 0 \tag{4.17}$$

and this is necessary to compute (4.11), with $\Pi(t)$ and $\alpha(t)$. In Eqs. (4.13) and (4.14), both $\Pi(t)$ and $\alpha(t)$ are actually independent of the dynamic characteristics of an observation mechanism, $h(t,\ x)$ and $R(t)$. Hence the optimal control depends on the cost rate function matrices M and N and on the system dynamics $f(t, x)$. However, a serious difficulty arises in the version of numerical computation on Eqs. (4.12), (4.13), (4.14), (4.15) and (4.16). In fact, the computation of (4.12) with Eqs. (4.13) and (4.15) has to start with the pre-assigned initial values of the state estimation $\hat{x}(t_0|t_0)$ and error covariance $P(t_0|t_0)$ and, furthermore, with $\Pi(t_0)$ and $\alpha(t_0)$ which are determined by the so-called trial and error method.

5. An illustrative example

For the purpose of exploring the quantitative aspects, we shall consider here the one-dimensional case. The dynamical system considered here is schematically shown by block diagram in Fig. 1. From Fig. 1, the stochastic differential equation of the dynamical system is given by

$$dx = f(-x)dt + udt + gdw \tag{5.1}$$

with

$$f(x) = \sin x. \tag{5.2}$$

The observation process is

$$dy = xdt + rdv. \tag{5.3}$$

Application of (3.4a) and (3.4b) to the present case gives

$$a(t) = -\sin\hat{x}\,\exp(-0.5p) \tag{5.4}$$

$$b(t) = -\cos\hat{x}\,\exp(-0.5p). \tag{5.5}$$

From Eqs. (4.4) and (3.26b), the approximated filter dynamics and related error covariance are determined by

$$d\hat{x} = -\sin\hat{x}\,\exp(-0.5p)dt - udt + pr^{-2}(dy - \hat{x}dt) \tag{5.6}$$

and

$$\frac{dp}{dt} = -2p\cos\hat{x}\,\exp(-0.5p) + g^2 - p^2r^{-2}. \tag{5.7}$$

Letting n = 1 and m = 1, in (4.1), we have

$$\hat{\psi}^{o}(t, \xi) = \pi(t)\xi + \alpha(t) \tag{5.8}$$

and

$$V(t, \xi) = \pi(t)\xi^2 + 2\alpha(t)\xi + \beta(t), \tag{5.9}$$

where

$$\frac{d\pi(t)}{dt} = \pi^2(t) - 1 \tag{5.10a}$$

$$\frac{d\alpha(t)}{dt} = \pi(t)\alpha(t) - \pi(t)a(t) \tag{5.10b}$$

$$\frac{d\beta(t)}{dt} = -\sigma^2(t)\pi(t) - 2a(t)\alpha(t) + \alpha^2(t) - p(t|t). \tag{5.10c}$$

Equations (5.6) to (5.10) are simulated on a digital computer with the subroutine for the generation of random disturbance, $\gamma(t)$ and $\theta(t)$. Figure 2(a) shows the running values of the state estimation $\hat{x}(t|t)$ (in figures presented here and below, the symbols $\hat{x}(t|t)$ and $p(t|t)$ are simply denoted by $\hat{x}(t)$ and $p(t)$), for the pre-assigned control interval [0, 1.0] (sec). The sample path behavior of the true system is also shown as the run $x(t)$. However, the $x(t)$-process is, in practice, in accessible and this is only for comparative observation. The dotted run in Fig. 2(a) shows the sample path behavior of the quasi-linearized system. Comparison of the sample paths of the quasi-linear system and filter dynamics with the that of true system, actully reveals that, as time goes on, the pursuit behavior of the $\hat{x}(t)$-process to the in-accessible $x(t)$-process becomes improved with the elevated accuracy of the stochastic linearization. The optimal control signal run is also plotted on Fig. 2(a). Figure 2(b) shows the error covariance of filtering action $p(t|t)$, and also $\pi(t)$, $\alpha(t)$ which may be adopted as a successful set of trial and error method. Figure 3 shows the averaged runs of ten sample paths.

6. Conclusion

The technique started with the stochastic linearization of the dynamical system and with that of the observation dynamics. Based on the linearized system dynamics, a class of finite dimensional approximations to the optimal filter has been drived. The optimal control has been obtained for the linearized system by means of solving Bellman's equation. In general, the optimal control depends parametrically on both the conditional averaged behavior $\hat{f}[t, x] = a(t)$ of nonlinear action and the choice of performance index factors M, N. Through the analytical development and the numerical results, we may conclude that the approximation procedure has desirable properties in realizing the feedback configuration of stochastic optimal control.

References

[1]. Joseph, P.D. and Tou, J.T.: On linear control theory, AIEE. Trans. Appl. Ind. Pt. II, 80, 193-196 (1961).

[2]. Schultz, P.R.: Optimal control in the presence of measurement errors and random disturbances, Ph. D. Thesis, Dept of Eng., Univ. of California, Los Angeles, California, (1963).

[3]. Wonham, W.M.: On the separation theorem of stochastic control; Tech. Reps. Control and Information System Lab., Elect. Res. Center, NASA, Rep. No. PM-38, Camb. Mass. Jan. (1968)

[4]. Itô, K: On stochastic differential equations; Mem. Math. Soc., 4, (1951).

[5]. Yaglom, A.M.: An introduction to the theory of random functions, Prentice-Hall, N.J. (1962).

[6]. Wong, E. and Zakai, M.: On the relation between ordinary and stochastic differential equations; Int. J. Eng. Sci. 3, (1965).

[7]. Doob, J.L.: Stochastic processes; John Wiley, N.Y. (1953).

[8]. Sunahara, Y.: An approximate method of state estimation for nonlinear dynamical systems; Center for Dynamical Systems, Tech. Reps.,

Div. of Appl. Math., Brown Univ., No. 67-8, (1967).

[9]. Kalman, R.E.: A new approach to linear filtering and prediction problems; Trans. ASME, J. Basic Eng., 82D, (1960), p.p. 33-45.

[10]. Kalman, R.E. and Bucy, R.S.: New results in linear filtering and prediction theory; Trans. ASME, J. Basic Eng., 83D, (1961), p.p. 95-108.

[11]. Dynkin, E.B.: Markov processes, I, II; Academic Press, N.Y., (1965)

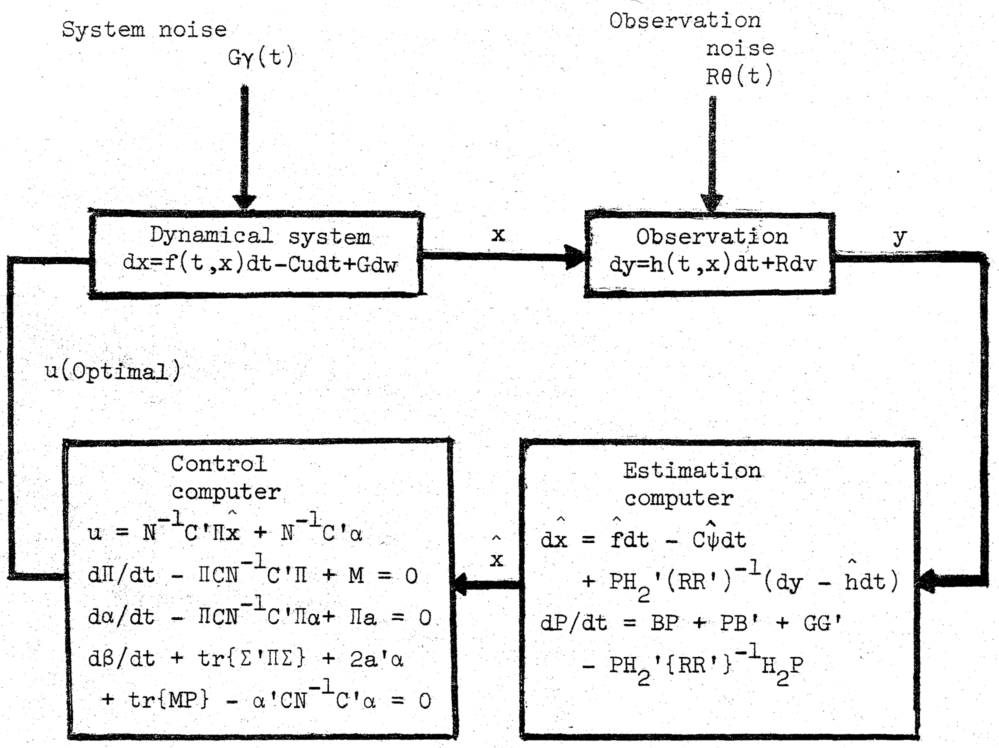


Fig. 1. Over-all configuration of optimal control for the nonlinear dynamical system under noisy observation

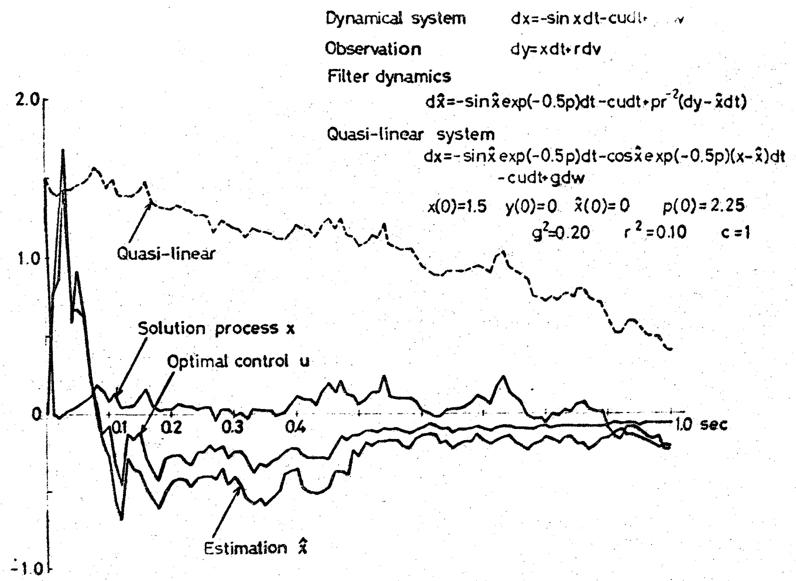
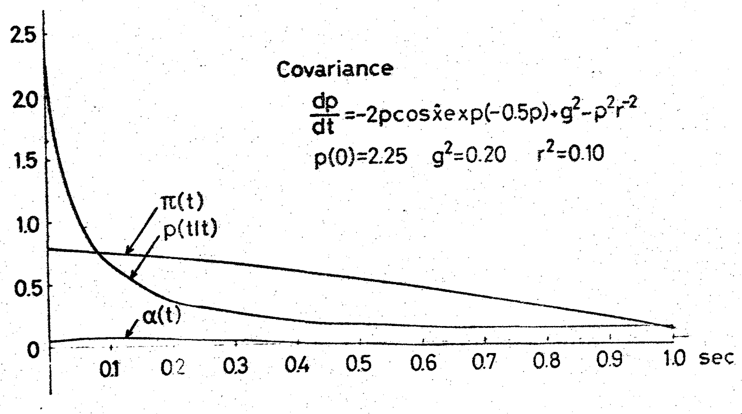


Fig. 2. (a) Sample path behaviors of the system, quasi-linearized system, filter and optimal control



(b) Error covariance $p(t|t)$ and convergence of $\pi(t)$ and $\alpha(t)$

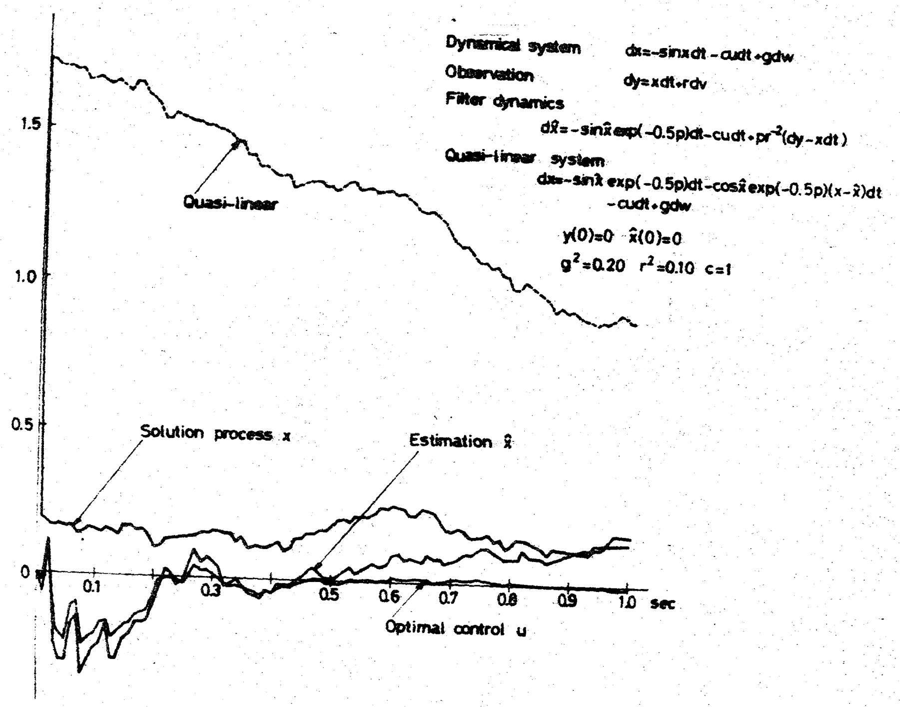


Fig. 3. Averaged runs of 10 sample paths